Research

Development and Evaluation of a Holistic and Mechanistic Modeling Framework for Chemical Emissions, Fate, Exposure, and Risk

Li Li,[1,2] Alessandro Sangion,[2,3] Frank Wania,[2] James M. Armitage,[4] Liisa Toose,[3] Lauren Hughes,[3] and Jon A. Arnot[2,3,5]

[1]School of Public Health, University of Nevada, Reno, Reno, Nevada, USA
[2]Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada
[3]ARC Arnot Research and Consulting, Toronto, Ontario, Canada
[4]AES Armitage Environmental Sciences, Inc., Ottawa, Ontario, Canada
[5]Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada

BACKGROUND: Large numbers of chemicals require evaluation to determine if their production and use pose potential risks to ecological and human health. For most chemicals, the inadequacy and uncertainty of chemical-specific data severely limit the application of exposure- and risk-based methods for screening-level assessments, priority setting, and effective management.

OBJECTIVE: We developed and evaluated a holistic, mechanistic modeling framework for ecological and human health assessments to support the safe and sustainable production, use, and disposal of organic chemicals.

METHODS: We consolidated various models for simulating the PROduction-To-EXposure (PROTEX) continuum with empirical data sets and models for predicting chemical property and use function information to enable high-throughput (HT) exposure and risk estimation. The new PROTEX-HT framework calculates exposure and risk by integrating mechanistic computational modules describing chemical behavior and fate in the socioeconomic system (i.e., life cycle emissions), natural and indoor environments, various ecological receptors, and humans. PROTEX-HT requires only molecular structure and chemical tonnage (i.e., annual production or consumption volume) as input information. We evaluated the PROTEX-HT framework using 95 organic chemicals commercialized in the United States and demonstrated its application in various exposure and risk assessment contexts.

RESULTS: Seventy-nine percent and 97% of the PROTEX-HT human exposure predictions were within one and two orders of magnitude, respectively, of independent human exposure estimates inferred from biomonitoring data. PROTEX-HT supported screening and ranking chemicals based on various exposure and risk metrics, setting chemical-specific maximum allowable tonnage based on user-defined toxicological thresholds, and identifying the most relevant emission sources, environmental media, and exposure routes of concern in the PROTEX continuum. The case study shows that high chemical tonnage did not necessarily result in high exposure or health risks.

CONCLUSION: Requiring only two chemical-specific pieces of information, PROTEX-HT enables efficient screening-level evaluations of existing and premanufacture chemicals in various exposure- and risk-based contexts. https://doi.org/10.1289/EHP9372

Introduction

More than 350,000 chemicals and mixtures have been registered in national and regional chemical inventories (Wang et al. 2020), amounting to global annual chemical sales of about 3.7 trillion Euros (European Chemical Industrial Council 2021). Although chemical production and use bring significant socioeconomic benefits and value, some exposures may also pose unacceptable risks to humans and ecological receptors. Regulations such as the U.S. Lautenberg Chemical Safety Act (i.e., LCSA 2016) (U.S. Congress 2016), the Canadian Environmental Protection Act (i.e., CEPA 1999) (Government of Canada 1999), and the European Registration, Evaluation Authorisation and Restriction of Chemicals (REACH) regulation (EC 2007) seek to evaluate and manage chemicals to ensure their safe production, use, and disposal. The large number of chemicals in commerce necessitates screening and priority setting for those posing the highest impacts on the environment and human health. Models are critical to this task because it is not feasible to measure concentrations of tens of thousands of chemicals in the multimedia environment, in organisms, and in human tissues.

Exposure science encompasses analyses of the chemical life cycle, emission and mode of entry, fate and transport in various environments, exposure factors relating receptor behavior to multimedia contact, and external and internal exposures (NRC 2012; National Academies of Sciences, Engineering, and Medicine 2017). These components constitute a PROduction-To-EXposure (PROTEX) continuum. Models for exposure and risk estimation may comprise the entire continuum or integrate one or more components of this spectrum. In its entirety such a model would allow users to examine how exposure and risk respond to changes in chemical production or use patterns to inform chemical management, for example, product development, proposed new uses, risk mitigation. The European Union System for the Evaluation of Substances (EUSES) provided the first effort to characterize the PROTEX continuum given that it supports the quantification of chemical exposure and risks for humans and several ecological receptors (Vermeire et al. 1997, 2005). However, EUSES has limited coverage for environmental compartments, exposure routes, and ecological receptors, with a prominent limitation being the lack of capacity to characterize human exposure to chemicals used and released indoors (Undeman and McLachlan 2011; van de Meent et al. 2014). The USEtox model provides more complete insights into exposure and health impacts because it integrates multiroute exposures from indoor or near-consumer environments (i.e., near-field) and the ambient environment distant from consumers (i.e., far-field) (Rosenbaum et al. 2008, 2011); however, it has limited ecological coverage. The PROTEX model (Li et al. 2018a, 2018b) supports mechanistic, time-variant simulations of chemical emissions, fate, and concentrations in indoor, urban, and

rural environments, as well as exposures of humans and other organisms. However, PROTEX is data and resource intensive; its dynamic nature is designed for higher-tiered comprehensive assessments and is not ideal for high-throughput (HT) screening applications. Efforts to combine different models for HT exposure and risk estimation include the ExpoDat initiative (Shin et al. 2015) and the U.S. Environmental Protection Agency (EPA)'s Systematic Empirical Evaluation of Models (SEEM) as part of the ExpoCast program (Ring et al. 2019; Wambaugh et al. 2013, 2014). The SEEM3 framework integrates the results of several exposure models and calibrates the consensus predictions to human exposure estimates for 114 chemicals in the U.S. population inferred from biomonitoring data, whereby it provides HT human exposure rates for more than 500,000 chemicals (Ring et al. 2019). However, SEEM does not provide mechanistic insights into the relationships between individual components in the PROTEX continuum and how they are related to various physical, chemical, physiological, behavioral, and social factors. There is a need for a holistic, mechanistic modeling framework that accounts for aggregate exposure to humans and ecological receptors relating directly to production for HT assessments.

The general paucity of chemical input information often precludes the application of exposure- and risk-based models (Egeghy et al. 2011; Wetmore et al. 2012). New approach methodologies (NAMs) are being developed to address data gaps in estimates of emissions (Li and Wania 2016; Tao et al. 2018; van de Meent et al. 2020), hazard (Gocht et al. 2015; Judson et al. 2010; Tice et al. 2013), exposure (Shin et al. 2015; Wambaugh et al. 2019), and risk (National Academies of Sciences, Engineering, and Medicine 2017; Patlewicz et al. 2018; Wetmore et al. 2015). Notably, a key source of uncertainty in exposure and risk estimation is uncertainty in chemical use and emission rates (Arnot et al. 2012; Breivik et al. 2012; Ring et al. 2019; Shin et al. 2015). Tools are being developed to improve the mechanistic understanding of chemical fate in the socioeconomic system (i.e., the technosphere), comprising all activities throughout the chemical life cycle (e.g., production, industrial processes, use, waste disposal) (Li and Wania 2016; Li 2020b; van de Meent et al. 2020). Quantitative structure–activity relationship (QSAR), quantitative structure–property relationship (QSPR), and quantitative structure–use relationship (QSUR)—collectively referred to as QSXR—have also been advanced to provide reliable estimates of physicochemical properties, use function, toxicokinetics, and environmental end points for a wider spectrum of chemicals (Arnot et al. 2014; Brown et al. 2019; Mansouri et al. 2018; Papa et al. 2014, 2018; Phillips et al. 2017). Regulatory agencies are considering the incorporation of NAMs in formal decision-making contexts (ECHA 2016b; Health Canada 2021; Kavlock et al. 2018; U.S. EPA 2018). Developing and evaluating NAMs is critical to advancing chemical assessments and fostering confidence in their applications.

Herein we introduce and evaluate a holistic and mechanistic chemical exposure and risk estimation framework named PROTEX-HT that characterizes and quantifies a chemical's journey from production to ecological and human receptors. The new PROTEX-HT model consolidates modules for simulating chemical emissions and mode of release, fate, and transport in representative indoor and natural environments, food web bioaccumulation in aquatic and terrestrial organisms, and exposures and potential risks to a range of representative ecological receptors and humans. PROTEX-HT is parameterized here with QSXR, which facilitates the operation of the system providing a multitude of chemical evaluation and management data based on only two pieces of chemical information, that is, tonnage (production or consumption) and molecular structure [e.g., simplified molecular-input line-entry system

notation (Weininger 1988)]. The use of QSXR predictions enables HT screening-level evaluations of existing and premanufacture chemicals. In this work, we apply PROTEX-HT to 95 organic chemicals commercialized in the United States and evaluate its performance by comparing model predictions with environmental monitoring and biomonitoring data. We also showcase how PROTEX-HT can guide decision-making for various chemical management objectives.

Methods

Overview of PROTEX-HT

Figure 1 depicts how PROTEX-HT integrates chemical releases from multiple stages in the chemical life cycle, fate, and transport in multimedia environments and exposure to ecological receptors and humans through multiple routes. To realize this integrative framework, PROTEX-HT combines the substance flow analysis model Chemicals in Products–Comprehensive Anthropospheric Fate Estimation (CiP-CAFE) (Li and Wania 2016; Li 2020a), the natural environmental fate and exposure model [Risk Assessment, IDentification And Ranking (RAIDAR)] (Arnot and Mackay 2008), and the indoor fate and consumer exposure model RAIDAR-Indoor and Consumer Exposure (ICE) (Li et al. 2018c). Text S1, "Description of components in PROTEX-HT" in the Supplemental Material, presents detailed information on the structure, configuration, and rationale of each model.

Provided with user-supplied chemical tonnages (annual production or consumption volume in metric tons per year), CiP-CAFE (version 2.0) calculates chemical flows between main life cycle stages and waste disposal practices, as well as rates of emissions therefrom. Specifically, CiP-CAFE characterizes an archetypal supply chain: After synthesis ("production" in Figure 1), a chemical is distributed between up to five end-use applications (indicated by five arrows in Figure 1), designated by their distribution ratios, for potential further use in manufacturing formulations and preparations ("industrial processes"), producing products readily used in professional and residential settings ("instantaneous use"), and articles with long service lives ("in service"). Wastes are generated from these life cycle stages, and along with end-of-life waste, treated in waste disposal facilities (engineered landfill, dumping and simple landfill, wastewater treatment, and so on).

Emissions occurring outdoors (outdoor air, surface water, and soil) are assigned to RAIDAR (version 3.0), which calculates the fate and transport of chemicals in an archetypal temperate North American ecosystem (Figure 1). RAIDAR predicts chemical concentrations in different far-field environmental compartments (air, surface water, soil, and sediment) by mechanistically quantifying various advection, diffusion, and reaction processes. It also quantifies chemical bioaccumulation and concentrations in a broad range of representative ecological receptors (plankton, invertebrates, fish, birds, and mammals) and agricultural organisms (root and foliage vegetation, cows, pigs, chickens, and so on) using mechanistic toxicokinetic models. These agricultural organisms constitute food for the human population parameterized here with the anthropometric, physiological, dietary, and activity data representative of a North American male adult. RAIDAR calculates the daily human exposure rate [in nanograms chemical per kilogram body weight (BW) per day] through inhalation, consumption of drinking water, and dietary ingestion.

Emissions occurring indoors (indoor air and direct applications to the body) are assigned to RAIDAR-ICE (version 1.5), which calculates the fate and transport of chemicals in an archetypal North American home (Figure 1). RAIDAR-ICE predicts chemical concentrations in different near-field compartments (indoor air, foam furniture, carpet, flooring, hard surfaces, and the dust thereon, and
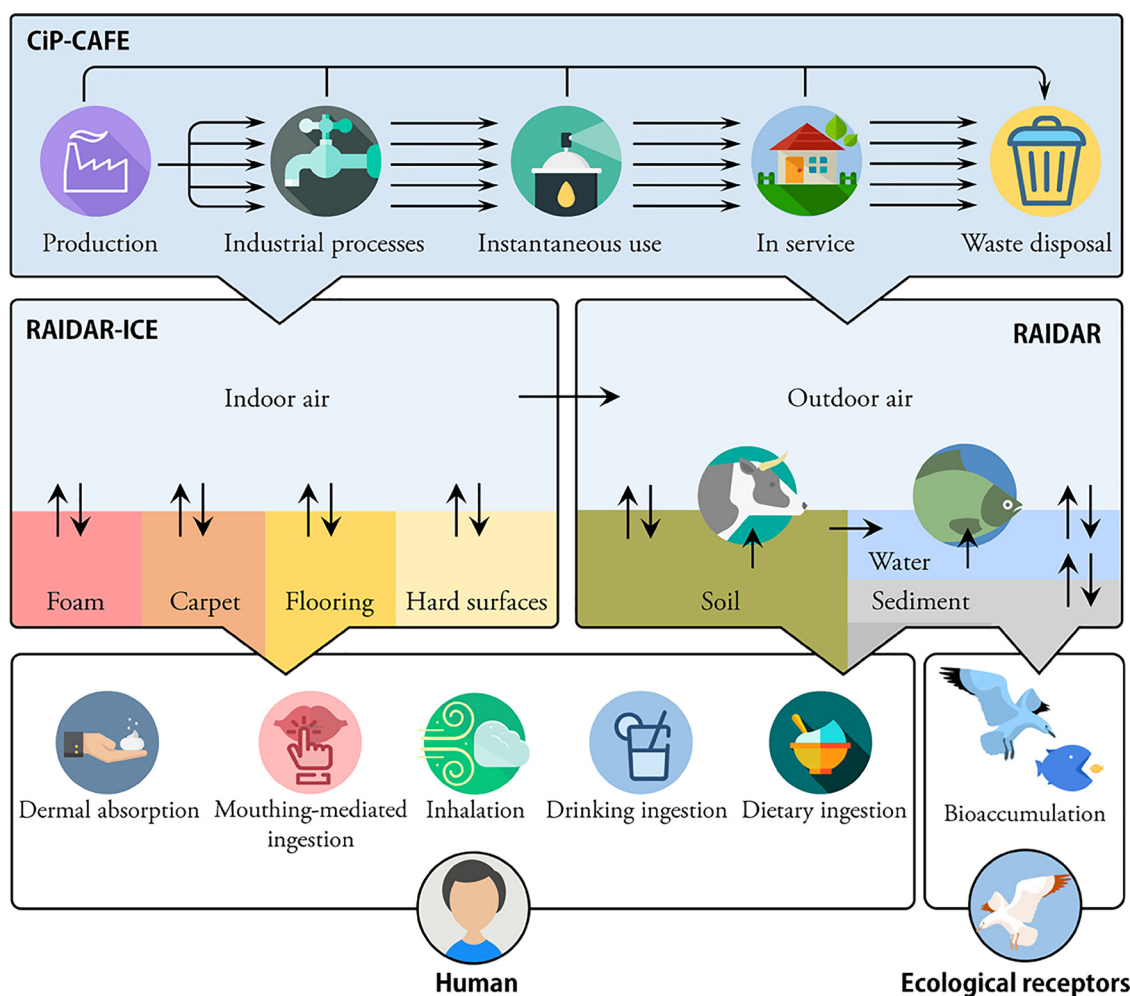
Figure 1. Schematic overview of PROTEX-HT and the three main modular components: CiP-CAFE [converting chemical tonnages to emission rates (Li and Wania 2016; Li 2020a)], RAIDAR and RAIDAR-ICE [converting emission rates to concentrations and exposures (Arnot and Mackay 2008; Li et al. 2018c)]. Also indicated are life cycle and waste treatment stages considered in CiP-CAFE and environmental and biological compartments. Arrows indicate some of the flow of chemicals. For details of the modules, see Text S1, "Description of components in PROTEX-HT" in the Supplemental Material. Note: CiP-CAFE, Chemicals in Products–Comprehensive Anthropospheric Fate Estimation; HT, high throughput; ICE, Indoor and Consumer Exposure; PROTEX, PROduction-To-EXposure; RAIDAR, Risk Assessment, IDentification And Ranking.

so on.) by mechanistically quantifying advection, diffusion, and reaction processes. The same representative human in RAIDAR is included in RAIDAR-ICE. RAIDAR-ICE calculates the daily exposure rate (in nanograms chemical per kilogram BW per day) through inhalation of indoor air, mouthing-mediated ingestion (i.e., ingestion of chemicals through the hand- and object-to-mouth contact) (Li et al. 2021), and dermal absorption. RAIDAR-ICE also calculates chemical fluxes ventilated from indoors to outdoors, which are added to the outdoor emissions for RAIDAR modeling (see Text S1, "Description of components in PROTEX-HT" in the Supplemental Material).

The predicted route-specific exposure rates by the two models can be aggregated to give the overall daily exposure rate. Both models include a one-compartment physiologically based model capturing key pharmacokinetic/toxicokinetic processes to convert the external daily exposure rate to internal doses, such as lipid-normalized whole-body concentrations, blood and urine concentrations. The prediction of biological concentrations provides opportunities for comparisons with biomonitoring data. RAIDAR and RAIDAR-ICE are steady-state models, requiring a time-invariant emission rate as input. We, therefore, parameterized CiP-CAFE with constant tonnages and ran the model for 100 y [a sufficiently long time for most long-lived articles (Li and Wania

2018)] to generate steady-state emission rates compatible with the RAIDAR and RAIDAR-ICE input requirement although the original version of CiP-CAFE supports predicting dynamic changes in chemical flows and emissions for up to 200 y if supplied with time-variant chemical tonnages.

Case Study with 95 Synthetic Organic Chemicals

We predicted the concentrations of 95 synthetic organic chemicals in various media of an archetypal environment in the United States as well as exposure rates to the general population. The selected chemicals belong to different functional use categories and include 68 biocides, 4 intermediate and/or raw chemicals, 4 construction material additives, 4 solvents, 8 plasticizers, and 7 personal care product ingredients. They also have diverse physicochemical, biological, and toxicological properties (Excel Table S1). The list contains several legacy persistent organic pollutants restricted or banned in the United States, such as organochlorine biocides (e.g., hexachlorobenzene, and α-, β-, and γ-hexachlorocyclohexanes). The list also contains four polycyclic aromatic hydrocarbons (PAHs) as construction material additives because they are synthesized industrially and used in sealing and coating products for construction, road, and pavement according to the Toxic Substances

Table 1. Key parameters required (indicated by "X") by the PROTEX-HT components: CiP-CAFE (Li and Wania 2016; Li 2020a), RAIDAR (Arnot and Mackay 2008), and RAIDAR-ICE (Li et al. 2018c).

Parameter	CiP-CAFE	RAIDAR	RAIDAR-ICE	QSXR predictions
Molar mass	X	X	X	—
Molar volume	X	—	—	ACD/Labs (Release 2019.2.1)
Equilibrium partition coefficients (K_{OW} and K_{OA})[a]	X	X	X	Consensus value (geometric mean) of predictions made with OPEn Structure-activity/property Relationship App (OPERA) (Mansouri et al. 2018), poly-parameter linear free energy relationships (ppLFERs) (Ulrich et al. 2017), and EPI Suite (U.S. EPA 2012)
Acidity/basicity and dissociation constants (pK_a and pK_b)	—	X	—	ACD/Labs (Release 2019.2.1)
Air hydroxylation and ozonation rate constants (k_{OH} and k_{O3})[b]	X	X	X	EPI Suite AOPWIN model (U.S. EPA 2012)
Primary biodegradation half-life (HL_{biodeg})				Regression equations (Arnot et al. 2005, 2012) against outputs from the EPI Suite BIOWIN models (U.S. EPA 2012)
Degradation half-life in surface water	X	X	—	Assumed to be equal to the primary biodegradation half-life, based on an empirical relationship (Fenner et al. 2005)
Degradation half-life in soil	—	X	—	Assumed to be twice the primary biodegradation half-life, based on an empirical relationship (Fenner et al. 2005)
Degradation half-life in sediment or waste stock	X	X	—	Assumed to be 10 times the primary biodegradation half-life, based on an empirical relationship (Fenner et al. 2005)
Whole-body biotransformation half-life in fish (HL_{fish})	—	X	—	Consensus value (geometric mean) of predictions made with OPERA (Mansouri et al. 2018), the Iterative Fragment Selection (IFS) QSAR (Brown et al. 2012), and the fish biotransformation models (Papa et al. 2014) implemented in QSARINS-Chem (Gramatica et al. 2014)
Whole-body biotransformation half-life in mammals (including humans) (HL_{mammal})	—	X	X	Consensus value (geometric mean) of predictions made with IFS QSAR (Arnot et al. 2014) and the human biotransformation models (Papa et al. 2018) implemented in QSARINS-Chem (Gramatica et al. 2014)
Functional use category				QSUR (Phillips et al. 2017), corrected based on market information documented in the U.S. EPA's Functional Use Database (Isaacs et al. 2016)
Distribution ratios	X	—	—	Parameterized automatically based on the functional use category (this paper)
Emission, waste, and decomposition factors	X	—	—	Parameterized automatically based on the functional use category (this paper)
Lifespan of articles[c]	X	—	—	Parameterized automatically based on the functional use category (this paper)

Note: —, not applicable; CiP-CAFE, Chemicals in Products–Comprehensive Anthropospheric Fate Estimation; HT, high throughput; ICE, Indoor and Consumer Exposure; PROTEX, PROduction-To-EXposure; QSAR, quantitative structure–activity relationship; QSARINS-Chem, quantitative structure–activity relationship–Insubria chemistry; QSUR, quantitative structure–use relationship; QSXR, collectively, QSPR, QSAR, and QSUR; Risk Assessment, IDentification And Ranking.

[a]Predictions are for the neutral form if a chemical is ionizable. RAIDAR automatically calculates the fractions of neutral and ionized forms from dissociation coefficients based on the Henderson-Hasselbalch equation, as well as partition coefficients (also known as distribution ratios) for combined neutral and ionized forms. K_{AW} is calculated as K_{OW}/K_{OA}, that is, following the thermodynamic triangle.

[b]Hydroxylation half-lives in the outdoor and indoor air are calculated with assumed hydroxyl radical concentration to be 9.7×10^5 molecules/cm^3 outdoors and 1.7×10^5 molecules/cm^3 indoors, respectively. Ozonation half-lives in the outdoor and indoor air are calculated with assumed ozone concentration to be 7×10^{11} molecules/cm^3 outdoors and 3.5×10^{11} molecules/cm^3 indoors, respectively. The overall half-lives in the indoor and outdoor air ($HL_{indoor\,air}$ and $HL_{outdoor\,air}$) combine the corresponding hydroxylation and ozonation half-lives.

[c]Required only for chemicals in articles (i.e., chemicals used in objects whose functions are determined mainly by their shapes, surfaces, and designs; for details, see Text S1 "Description of components in PROTEX-HT" in the Supplemental Material).

Control Act (TSCA)'s 2016 Chemical Data Reporting (CDR) (U.S. EPA 2020). However, their reported tonnages do not include the quantities generated unintentionally through combustion.

These chemicals were selected because their nationally representative exposure rates had previously been inferred from biomonitoring data in the National Health and Nutrition Examination Survey (NHANES) (Wambaugh et al. 2014) and hence were available for evaluating PROTEX-HT's performance. The evaluation was based on inferred exposure rates of the total population (males and females combined), in accordance with the practice in SEEM (Ring et al. 2019; Wambaugh et al. 2013, 2014); in fact, no significant differences were found between the total population and either males or females (Wambaugh et al. 2014). We compared the aggregate exposure rates predicted by PROTEX-HT (central-tendency estimates based on the central tendency of chemical tonnage) with the medians of the NHANES-inferred exposure rates. The performance of PROTEX-HT was quantified using the coefficient of determination (R^2; calculated with Microsoft Excel), as well as the discrepancy between PROTEX-HT–predicted and NHANES-inferred exposure rates.

In addition to exposure rates, PROTEX-HT predicted chemical concentrations in environmental compartments and ecological receptors, whole-body concentrations in humans, as well as health risks for the 95 chemicals. We analyzed the correlation between the exposure rate, whole-body concentration, health risk, and chemical tonnage using Spearman's rank-order correlation coefficient (Spearman's ρ) calculated with Microsoft Excel.

Automated Parameterization of PROTEX-HT

The three primary modules of PROTEX-HT require inputs of chemical-specific information on properties and use function. Table 1 gives a complete list of key parameters. These parameters were estimated from the molecular structure using state-of-the-art QSXRs if experimentally determined values were not available. Following the Organisation for Economic Co-operation and Development's Guidance Documents on the Validation of QSAR Models (OECD 2007), we selected QSXR models with a) clearly defined predicted end points that share the same meaning as

the parameters required by the PROTEX-HT components; *b*) unambiguous, reproducible model algorithms; *c*) domains of applicability covering the majority, if not all, of the case study chemicals; and *d*) transparent information on goodness-of-fit, robustness, and predictivity. Excel Table S1 tabulates the QSXR-derived parameters for the 95 case study chemicals.

Specifically, well-established QSARs and QSPRs were used to parameterize the partitioning and reactive properties (Table 1). The QSUR developed by Phillips et al. (2017) was used to predict a chemical's function, that is, its functional role in the end-use product and/or article applications. The predicted functional uses were checked individually with market information documented in the U.S. EPA's Functional Use Database (Isaacs et al. 2016) if available; corrections were made when misclassification arose. Each functional use was then matched with 1 of 87 functional use categories defined in CiP-CAFE (Excel Table S2). For instance, di-*n*-octyl phthalate [Chemical Abstracts Service Registry Number (CASRN) 117-84-0] was predicted to be a fragrance by the QSUR but corrected to being a plasticizer based on records in the U.S. EPA's Functional Use Database, which corresponded to CiP-CAFE's functional use category "plasticizer". Based on the assigned functional use category, CiP-CAFE automatically selected appropriate distribution ratios, by searching a built-in database documenting the relative likelihood that a function is found in different end-use applications (Breivik et al. 2012), to split the total chemical tonnage into up to five end-use applications. For example, 41%, 20%, 4%, and 35% of the total chemical tonnage of a plasticizer were allocated to polymer/plastic materials, textiles, electrical/electronic equipment, and miscellaneous use, respectively (Breivik et al. 2012). The distribution ratios of all 87 functional use categories are tabulated in Excel Table S2. Meanwhile, for each end-use application, CiP-CAFE populated the emission, waste, and decomposition factors with data either collected from official documents or computed from physicochemical properties by modules built into CiP-CAFE. For instance, factors for "production," "industrial processes," and "instantaneous use" were taken from the SPecific Environmental Release Categories (SPERC) (Sättler et al. 2012), REACH's Environmental Release Categories (ERCs) (ECHA 2016a), and the European Union's Technical Guidance Documents on Risk Assessment (De Bruijn et al. 2002) (ordered from highest to lowest priority); factors for emissions from "instantaneous use" and "in service" stages were computed in the EmissionRate module; factors for "engineered landfilling" and "dumping and simple landfilling" were computed in the Model for Organic Chemicals in LAndfills (MOCLA) module, and those for "wastewater treatment" were computed in the SimpleTreat module. In addition, for articles in each end-use application, CiP-CAFE populated the lifespans of articles, ranging from 2 to 16.8 y, with typical data from the Lifespan database for Vehicles, Equipment, and Structures (LiVES) (Murakami et al. 2010).

Text S3, "Applicability domains of QSXR models" in the Supplemental Material, details the applicability domains of these QSXRs. Excel Table S3 contains information on which of the 95 case study chemicals fall within the applicability domains of these QSXRs. The 95 chemicals were all located in the applicability domains of at least one QSXR, and more than 80% of the chemicals fell within the applicability domains of most QSXRs (Excel Table S3).

For risk estimation, we also used the conditional toxicity value QSAR to predict the reference dose, which characterized the threshold of daily chemical uptake beyond which exposure can cause observable adverse effects on the function of an animal's whole body (i.e., the systemic toxicity) (Wignall et al. 2018). PROTEX-HT characterizes the health risk using a risk assessment factor, defined as a unitless ratio between the predicted exposure rate and the reference dose.

Compilation of Chemical Tonnage Data

National chemical tonnages were assembled and curated from publicly available sources, including the 2016 CDR (given the highest priority if available), the Estimated Annual Agricultural Pesticide Use for Counties of the Conterminous United States, the Crop Protection Research Institute data set, and the U.S. EPA's High Production Volume list. Chemicals restricted or banned in the United States (chlorpyrifos; hexachlorobenzene; α-, β-, and γ-hexachlorocyclohexanes; nitrofen; and pentachlorophenol) were assumed to be associated with a tonnage of between 1 and 11.34 metric tons per year; the latter is the minimum reporting threshold by the 2016 CDR. For details of data compilation and hyperlinks to these data sources, see Text S2, "Compilation of chemical tonnage data" in the Supplemental Material. Excel Table S1 documents the assembled and curated data for PROTEX-HT modeling. In a generic case, we assumed that 10% of the chemical tonnage is produced and used in the modeled region (the default 10% rule in risk assessments) (De Bruijn et al. 2002). This is also consistent with assuming that 10% of the U.S. population lives in the modeled region (see Text S1, "Description of components in PROTEX-HT" in the Supplemental Material). All chemical tonnages were expressed as bins (ranges): The upper and lower bounds were used to generate the high and low estimates of exposure in this work, whereas their geometric means (central-tendency estimates) were used for chemical comparisons and ranking.

Uncertainty Analysis

We analyzed the overall uncertainty associated with the modeled exposure rate to illustrate how PROTEX-HT predictions were impacted by the propagation of uncertainties in the quantitative QSPR/QSAR predictions and the qualitative (classification) QSUR predictions. First, we assessed the magnitude of variation in the modeled exposure rate (U; in percentage of central-tendency estimate) caused by the propagation of inherent uncertainties in the QSPR- and QSAR-predicted partitioning and reactive properties (U_{Ii}), using the method proposed by MacLeod et al. (2002),

$$U = \sqrt{\sum_i (S_{Ii}^2 \times U_{Ii}^2)} \qquad (1)$$

where S_{Ii} is the sensitivity of the exposure rate (model output) to each QSPR and QSAR prediction I_i (model input), calculated by the percentage change in model output normalized by a 20% change in each model input (i.e., each input parameter was increased and decreased by 10%). Here, the sensitivity analysis is done for a hypothetical chemical with a molar mass of 250 g/mol, log K_{OA} of 9, log k_{OW} of 4, k_{OH} of 2×10^{-11} cm^3/(molecules · s), HL_{biodeg} of 200 h, HL_{fish} of 20 h, and HL_{mammal} of 15 h, since these numbers represent medians of properties of the 95 chemicals investigated here. For each QSPR or QSAR model, we defined the uncertainty U_{Ii} as the extent to which its predictions deviate from the experimentally determined values in its validation set, which was quantified using the root-mean-square-error (RMSE) or standard deviation in its external validation (see data in Text S4, "Uncertainties in quantitative QSPR and QSAR predictions" in the Supplemental Material). On the other hand, we performed an additional set of PROTEX-HT simulations, using the uncorrected QSUR predictions of chemical functions, to assess the impact of QSUR's misclassification on the
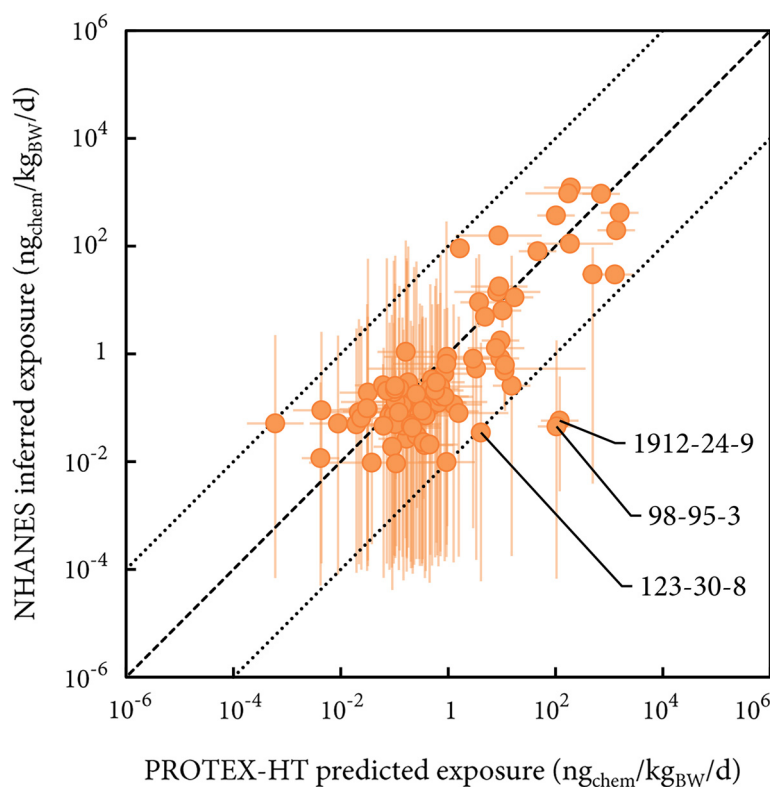
Figure 2. Comparison of aggregate exposure rates predicted by PROTEX-HT (central-tendency estimates based on the central tendency of chemical tonnage, and ranges derived from the high and low estimates of chemical tonnage) and inferred by Wambaugh et al. (2014) from human biomonitoring data in the U.S. National Health and Nutrition Examination Survey (NHANES) (medians and 95% confidence intervals). The dashed diagonal line represents perfect agreement between predictions and inferred values; the dotted lines represent a difference of two orders of magnitude. Chemicals with a difference of greater than two orders of magnitude are identified by Chemical Abstracts Service Registry Number. For the numerical values of the PROTEX-HT predictions, see Excel Table S4. Note: BW, body weight; HT, high throughput; PROTEX, PROduction-To-EXposure.

modeled exposure rate. We compared the predicted exposure rates with and without corrections for the misclassification.

Results

Evaluation of Model Performance

Figure 2 presents the PROTEX-HT–predicted aggregate exposure rates for the 95 case study chemicals. The predicted exposure rates spanned more than six orders of magnitude, from 6.0×10^{-4} (*p*-nitroanisole; CASRN 100-17-4) to 1.6×10^{3} (benzyl butyl phthalate; CASRN 85-68-7) nanograms chemical per kilogram BW per day. These predictions were also compared with exposure rates inferred from biomonitoring data in the NHANES (Wambaugh et al. 2014). Figure 2 indicates that our predictions were in satisfactory agreement with the NHANES-inferred exposure rates, with a coefficient of determination (R^2) of 0.59, which means PROTEX-HT (central-tendency estimates based on the central tendency of chemical tonnage) explained 59% of the variance observed in the medians of NHANES-inferred exposure rates. PROTEX-HT reproduced the medians of NHANES-inferred exposure rates for 75 chemicals (79% of the 95 chemicals) with a discrepancy smaller than an order of magnitude and for 92 chemicals (97%) with a discrepancy smaller than two orders of magnitude. Furthermore, the discrepancy between predicted and NHANES-inferred exposure rates was largest for the 4 intermediate and/or raw chemicals (with an RMSE of 2.2 log units) but similar for the remaining functional use categories (RMSEs of 1.12 log units for solvents, 0.82 log units for biocides, 0.73 log units for plasticizers, 0.71 log units for construction material additives, and 0.61 log units for personal care product ingredients). That is, the

model's performance was not skewed toward certain functional use categories. Nevertheless, PROTEX-HT substantially over-predicted exposure to nitrobenzene (CASRN: 98-95-3; 2,320 times), atrazine (CASRN: 1912-24-9; 2,080 times), and 4-amino-phenol (CASRN: 123-30-8; 117 times) (Figure 2).

The performance of PROTEX-HT was similar to that of the SEEM3 framework, which is an empirical machine learning approach that predicts exposure rates using a consensus Bayesian regression combining multiple model predictions (Ring et al. 2019; Wambaugh et al. 2014). For comparison, Figure S1 shows the performance of SEEM3 in predicting exposure rates for the 95 chemicals investigated here. SEEM3 succeeded in explaining 58% of the variations observed in the NHANES-inferred exposure rates and in predicting exposure rates for 76 and 89 chemicals with a difference within an order and two orders of magnitude, respectively. Here, PROTEX-HT predictions were made independently of NHANES inferences; by contrast, the SEEM3 exposure estimates had been a result of calibration to NHANES inferences and were not a true external evaluation of the framework. Interestingly, SEEM3 tended to substantially underestimate (>500 times) human exposure to persistent legacy biocides that have long been phased out, such as α-, β-, and γ-hexachlorocyclohexanes (CASRN 608-73-1, CASRN 319-85-7, and CASRN 58-89-9) and pentachlorophenol (CASRN 87-86-5). This discrepancy had been hypothetically attributed to the uncertainty associated with the use of production volumes instead of actual emission rates in SEEM3 (Ring et al. 2019) because persistence makes these biocides maintain disproportionately high levels in the environment relative to their production volumes. This process was captured in the mechanistic PROTEX-HT framework, which gave predictions in satisfactory (~5 to 10 times)

Table 2. Uncertainties associated with the modeled exposure rates (in percentage of central-tendency estimate) via exposure routes from the far-field environment, as a result of the propagation of uncertainties in partitioning (equilibrium partition coefficients) and reactive properties (air hydroxylation rate constant, primary biodegradation half-life, and whole-body biotransformation half-lives in fish and mammals) predicted by the QSARs and QSPRs adopted by PROTEX-HT, arrayed by assumed mode of entry.

Mode of entry	Input parameter	Inhalation of outdoor air (%)	Drinking ingestion (%)	Dietary ingestion (%)
100% to air	Overall	58	418	352
	K_{OA}	27	60	179
	K_{OW}	3	10	143
	k_{OH}	27	27	27
	HL_{biodeg}	0	321	0
	HL_{fish}	0	0	0
	HL_{mammal}	0	0	3
100% to water	Overall	1,106	326	596
	K_{OA}	750	0	2
	K_{OW}	13	10	175
	k_{OH}	27	0	0
	HL_{biodeg}	317	316	316
	HL_{fish}	0	0	102
	HL_{mammal}	0	0	0
100% to soil	Overall	1,144	1,257	1,048
	K_{OA}	248	1	4
	K_{OW}	548	529	717
	k_{OH}	27	0	1
	HL_{biodeg}	322	727	326
	HL_{fish}	0	0	0
	HL_{mammal}	0	0	0

Note: HT, high throughput; PROTEX, PROduction-To-EXposure; QSAR, quantitative structure–activity relationship; QSPR, quantitative structure–property relationship.

Table 3. Uncertainties associated with the modeled exposure rates (in percentage of central-tendency estimate) via exposure routes from the near-field environment, as a result of the propagation of uncertainties in partitioning (equilibrium partition coefficients) and reactive properties (air hydroxylation rate constant, primary biodegradation half-life, and whole-body biotransformation half-lives in fish and mammals) predicted by the QSARs and QSPRs adopted by PROTEX-HT, arrayed by assumed mode of entry.

Mode of entry	Input parameter	Inhalation of indoor air (%)	Dermal absorption (%)	Mouthing-mediated ingestion (%)
100% to air	Overall	67	889	2,854
	K_{OA}	35	738	1,501
	K_{OW}	28	141	1,339
	k_{OH}	5	10	14
	HL_{biodeg}	0	0	0
	HL_{fish}	0	0	0
	HL_{mammal}	0	0	0
100% to human skin	Overall	121	708	913
	K_{OA}	70	405	483
	K_{OW}	46	301	428
	k_{OH}	4	2	2
	HL_{biodeg}	0	0	0
	HL_{fish}	0	0	0
	HL_{mammal}	0	0	0

Note: HT, high throughput; PROTEX, PROduction-To-EXposure; QSAR, quantitative structure–activity relationship; QSPR, quantitative structure–property relationship.

agreement with NHANES inferences, given that PROTEX-HT's steady-state assumption considered the accumulation of chemical mass in the environment.

The most prominent advantage of PROTEX-HT is that it can simulate chemical behavior and fate in a stepwise manner throughout the PROTEX continuum. For instance, PROTEX-HT–predicted concentrations in various environmental media (e.g., indoor dust, indoor air, outdoor air) and ecological receptors (e.g., fish, beef cattle) can be compared with environmental and food monitoring data as a means of evaluating modeling performance (Table S1). Overall, for most chemicals, PROTEX-HT predictions fell well into the measured concentration ranges gathered from published studies (Table S1). An expected exception is bisphenol-A (raw material for polycarbonate plastic products), for which PROTEX-HT predictions were orders of magnitude lower than measurements, notably in the indoor environment (Table S1), primarily because PROTEX-HT assumed that raw materials were subject to complete reaction and do not appear in consumer products. PROTEX-HT also slightly underpredicted the concentrations of PAHs, primarily because unintentional combustion sources were omitted.

Despite PROTEX-HT's encouraging performance, uncertainties in model predictions need to be acknowledged. Tables 2 and 3 show that the propagation of uncertainties inherent in the QSPRs and QSARs used to estimate partitioning and reactive properties generally caused an overall uncertainty of an order of magnitude (i.e., <1,000%) in the modeled rates of human exposure, which is smaller than the 95% confidence intervals of NHANES-inferred exposure rates of all the 95 chemicals (∼6 orders of magnitude; vertical error bars in Figure 2). As such, the overall uncertainty did not prevent PROTEX-HT from distinguishing exposure rates between chemicals. However, when chemicals were released predominantly to indoor air, the estimated rates of mouthing-mediated ingestion were highly uncertain (up to a factor of ∼30). Compared with reaction half-lives,

uncertainties in partition coefficients made a greater contribution to the overall uncertainty in exposure estimates. On the other hand, the QSUR model used in this work misclassified the functional use categories of certain chemicals, leading to the use of inappropriate emission, waste, and decomposition factors, lifespans of articles, and distribution ratios and thus causing biases in exposure estimates. As Figure 3 shows, the misclassification changed the predicted exposure rates by a factor of >5 for 37 chemicals. Notably, the misclassification led to a severe overestimation of the exposure rates of nicosulfuron (CASRN 111991-09-4) and sulfosulfuron (CASRN 141776-32-1), by factors of 368 and 333, respectively. These two biocides were misidentified as personal care product ingredients. By contrast, the misclassification exerted limited impacts on the exposure rates of other chemicals, for example, pyrene (CASRN 129-00-0; with a difference of a factor of 2), which is a construction material additive misidentified as a colorant in the QSUR prediction.

Application 1: Ranking and Prioritizing Chemicals

PROTEX-HT supported the evaluation of chemicals based on a diverse array of metrics to meet requirements in various exposure- and risk-based assessment contexts. To illustrate some of the opportunities that PROTEX-HT provides, we continued to use the 95 case study chemicals. Figure 4 ranks these chemicals based on predicted exposure rate (in nanograms chemical per kilogram BW per day; Figure 4a), whole-body concentration (in nanograms chemical per gram lipid; Figure 4b), a risk assessment factor (RAF, unitless; Figure 4c), and regional chemical tonnage (in metric tons per year; Figure 4d). Spearman's ρ quantifies the similarity between rankings (Table S2): A higher Spearman's ρ indicates that two rankings are more consistent. An overall conclusion drawn from the inspection of Figure 4 is that rankings and subsequent priority setting were sensitive to the evaluation metrics being considered.

A comparison between Figure 4a and 4b indicates that the exposure rate was an acceptable surrogate for whole-body concentration in human risk assessments, with a Spearman's ρ = 0.88 (Table S2), at least for these 95 chemicals. Compared with the predicted exposure rate (Figure 4a), the predicted whole-body concentration had a wider range, spanning over seven orders of magnitude (Figure 4b).
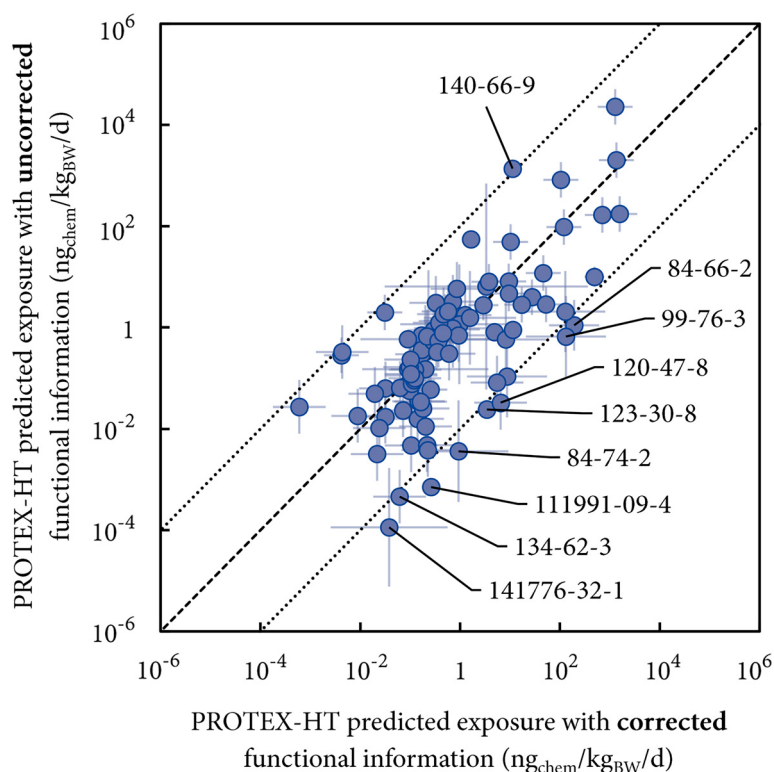
Figure 3. Comparison of aggregate exposure rates predicted by PROTEX-HT (central-tendency estimates based on the central tendency of chemical tonnage, and ranges derived from the high and low estimates of chemical tonnage) with corrected and uncorrected QSUR-derived functional information. The dashed diagonal line represents perfect agreement between two sets of predictions; the dotted lines represent a difference of two orders of magnitude. Chemicals with a difference of greater than two orders of magnitude are identified by Chemical Abstracts Service Registry Number. For corrected and uncorrected functional information, see Excel Table S1. Note: BW, body weight; HT, high throughput; PROTEX, PROduction-To-EXposure; QSUR, quantitative structure–use relationship.

The difference between the rankings based on exposure rate and whole-body concentration was the most remarkable for recalcitrant, hydrophobic chemicals. A prominent example is hexachlorobenzene (CASRN 118-74-1), ranked second for whole-body concentration (Figure 4b), despite its low exposure rate (ranked 31 in Figure 4a). Hexachlorobenzene's persistence prevents extensive biotransformation (whole-body biotransformation half-life of 2.5×10^4 h for a 70-kg person; Excel Table S1), and its hydrophobicity (log $K_{OW} = 5.73$; Excel Table S1) prevents efficient excretion through fecal egestion and urination (Li et al. 2019; Zhang et al. 2021). This was also the case for other persistent organic pollutants, such as α-, β-, and γ-hexachlorocyclohexane (ranked 12 to 14 for whole-body concentration vs. 39 to 43 for exposure rate). By contrast, rankings based on whole-body concentration and exposure rate were closer for chemicals subject to relatively fast elimination, such as phthalates (plasticizers) and parabens (chemicals in personal care products).

Figure 4c shows that the calculated RAFs were all <1, indicating no unacceptable health impacts posed on the general U.S. population based on the assessment end point selected here. The difference between rankings in Figure 4a and 4c, with a Spearman's $\rho = 0.59$ (Table S2), largely reflects the variation in reference dose. More potent chemicals may possess a high ranking even if the exposure rate is relatively low. For instance, ethyl *p*-nitrophenyl phenylphosphorothioate (EPN; CASRN 2104-64-5) had the lowest reference dose (highest toxicity) of 0.00001 milligrams chemical per kilogram BW per day among the investigated chemicals (Excel Table S1), which elevated its ranking from 68 for the exposure rate to 7 for RAF. On the other hand, a chemical with a high exposure rate may not necessarily have a high RAF. For instance, dibutyl phthalate (CASRN 84-74-2) had a relatively high reference dose (low toxicity) of 0.1 milligrams chemical per kilogram BW per day; it was ranked 80 for RAF despite a relatively high ranking of 34 for exposure rate.

Application 2: Setting Chemical-Specific Maximum Allowable Tonnage for Tailored Risk Estimation

The CiP-CAFE module in PROTEX-HT calculated an emission-to-tonnage ratio, namely, the percentage of the regional chemical tonnage (geometric mean) entering the environment (Figure 5a). Overall, the emission-to-tonnage ratio was rather low for chemicals used in industries, such as intermediates and/or raw chemicals ($\sim 5\%$), and high for chemicals with nonpoint open use, such as biocides and personal care chemicals ($\sim 50–100\%$). This ratio is often related to measures of emission control: Chemicals used in industries are often subject to end-of-pipe treatments to reduce or minimize emissions, whereas evaporative emissions of biocides are largely passive and difficult to control. Furthermore, the emission-to-tonnage ratio varied substantially between plasticizers, mostly depending on their propensity for evaporation to air: di-*n*-octyl phthalate (CASRN 117-84-0), di(2-ethylhexyl) phthalate (CASRN 117-81-7), and dicyclohexyl phthalate (CASRN 84-61-7) had low emission-to-tonnage ratios because of their high K_{OA} ($>10^{11}$), whereas the rest of the plasticizers had high emission-to-tonnage ratios ($>80\%$).

Likewise, PROTEX-HT calculated an exposure-to-tonnage ratio (Figure 5b), that is, the fraction of the regional chemical tonnage (geometric mean) taken in by the entire population living in the modeled region (expressed in parts per million because the numbers are small). The exposure-to-tonnage ratio is conceptually
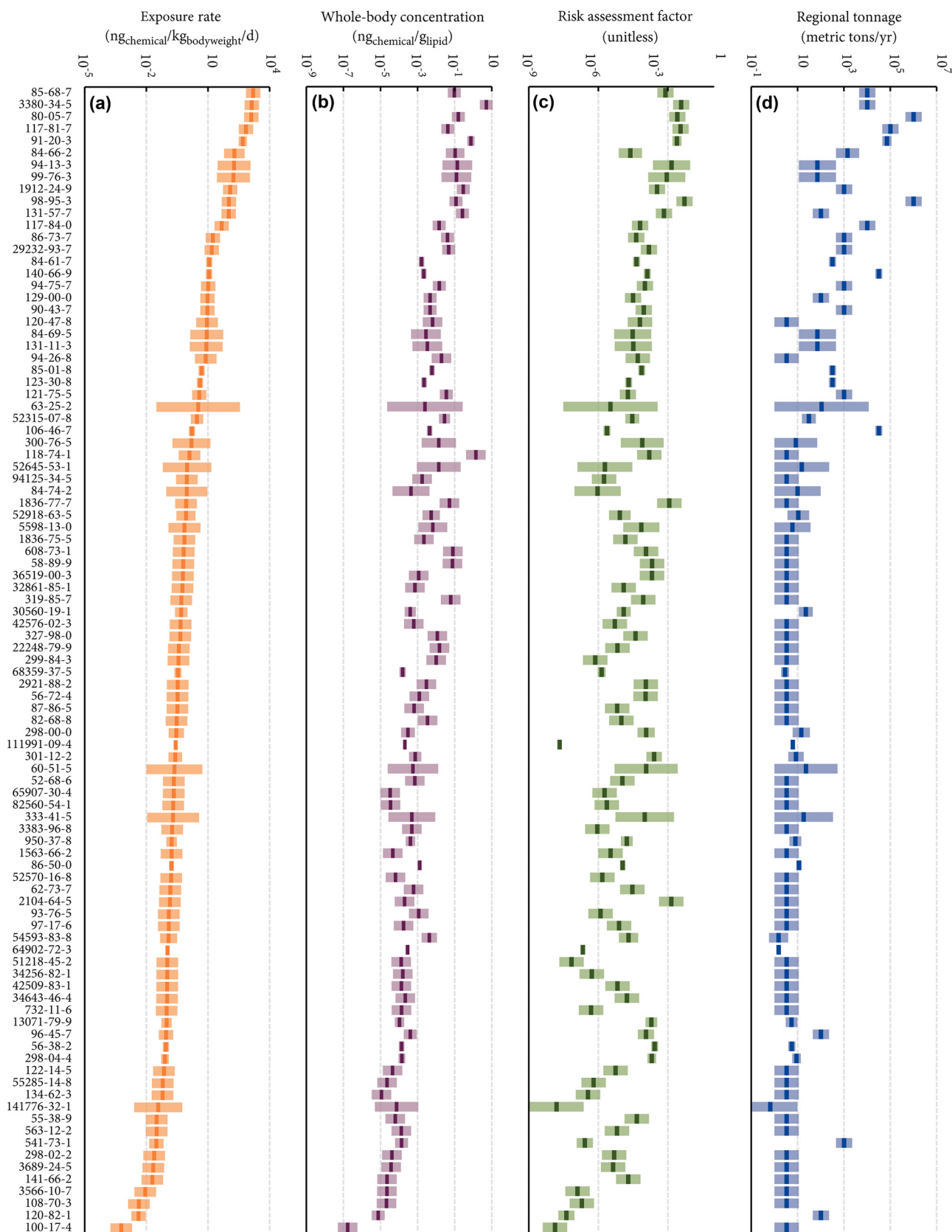
Figure 4. PROTEX-HT predicted (a) exposure rates, (b) whole-body concentrations, (c) risk assessment factors, and (d) regional tonnages of 95 synthetic organic chemicals (identified by their Chemical Abstracts Service Registry Numbers) in the United States, ranked by the geometric means of exposure rates. For the numerical values of the PROTEX-HT predictions, see Excel Tables S1 and S4. Note: HT, high throughput; PROTEX, PROduction-To-EXposure.
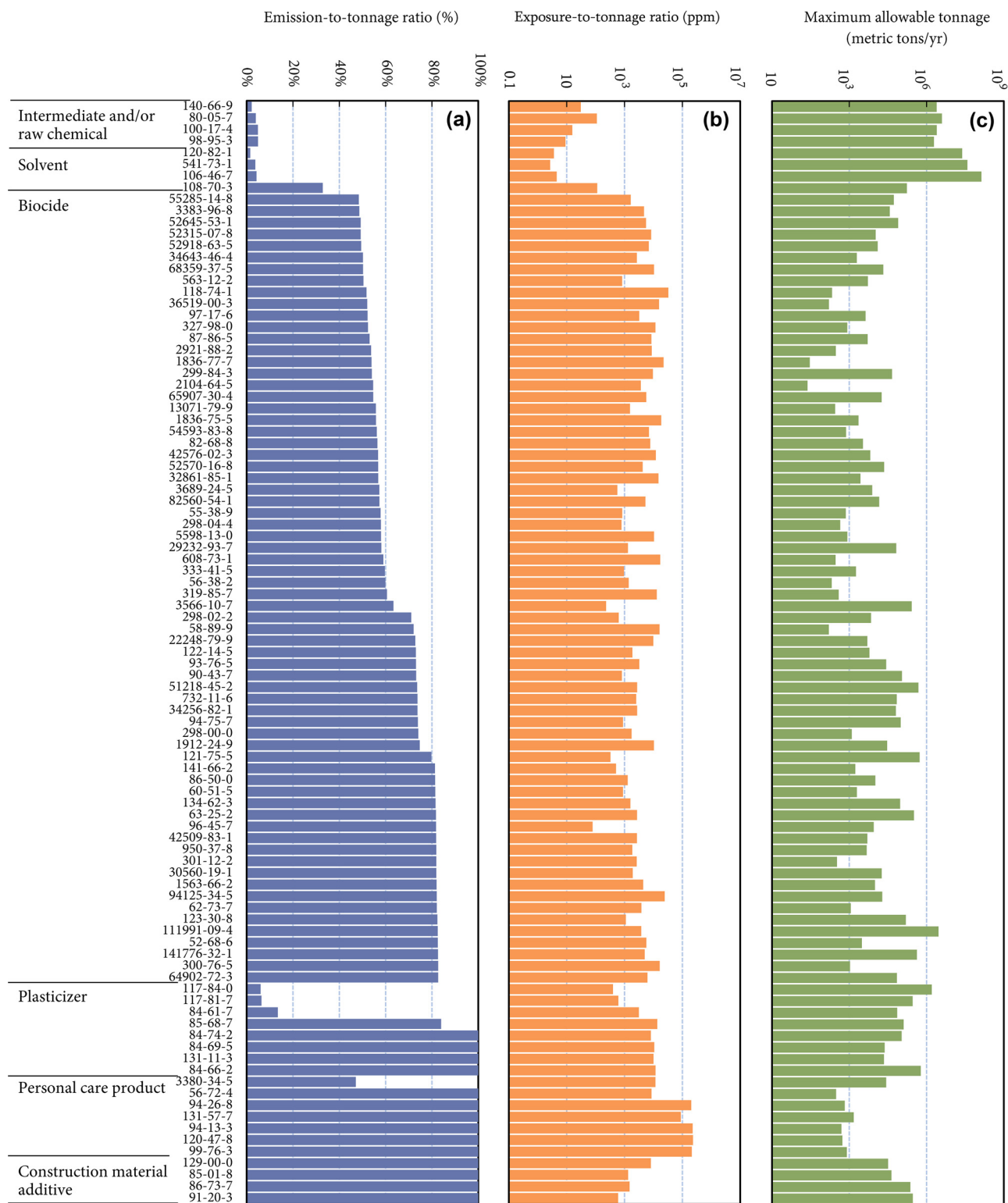
Figure 5. PROTEX-HT–predicted (a) emission-to-tonnage ratios, (b) exposure-to-tonnage ratios, and (c) maximum allowable tonnages of 95 synthetic organic chemicals (identified by their Chemical Abstracts Service Registry Numbers) in the United States, categorized by functional category and ranked by the emission-to-tonnage ratio. For the numerical values of the PROTEX-HT predictions, see Excel Table S5. Note: HT, high throughput; PROTEX, PROduction-To-EXposure.

similar to the intake-to-production ratio previously defined and calculated by Nazaroff et al. (2012) based on biomonitoring and surveyed manufacturing data. Table S3 shows that for nine chemicals discussed by Nazaroff et al. (2012), the two approaches gave fairly consistent estimates, with discrepancies of generally less than a factor of 10. As shown in Figure 5b, the exposure-to-tonnage ratio

was low for intermediates and/or raw chemicals ($\sim$10–100 ppm), moderate for biocides ($\sim$100–33,000 ppm) and plasticizers ($\sim$400–13,000 ppm), and high for chemicals with direct application onto the human skin, such as chemicals in personal care products ($\sim$6,000–170,000 ppm). This variability is partially explained by functional use categories having different emission-

to-tonnage ratios (Figure 5a). Another reason is that chemicals in different use categories may differ in their potential to enter the body. For example, chemicals used indoors typically have notably higher intake fractions than chemicals released outdoors (Zhang et al. 2014). In addition, within each functional use category, the exposure-to-tonnage ratio varied (Figure 5b) because of variability in chemical properties, such as the potential for accumulation in the environment and food webs.

In addition to a forward calculation mode that estimates risks based on the chemical tonnage, PROTEX-HT can back-calculate a maximum allowable tonnage for a specific chemical, that is, the critical quantity beyond which its production and use would cause unacceptable health risks, for example, based on the reference dose. Such a backward calculation sets a practical cap in chemical management, such that regulatory agencies or authorities can compare the cap and actual tonnage to estimate the margin of safety of chemical production and use in a jurisdiction. The back-calculated maximum allowable tonnages for the 95 chemicals in the modeled region (Figure 5c) spanned seven orders of magnitude, from 24 metric tons/y (EPN; CASRN 2104-64-5) to 135 million metric tons/y (1,4-dichlorobenzene; CASRN 106-46-7). In general, high maximum allowable tonnages were found for *a*) chemicals with a small emission-to-tonnage ratio, such as intermediate and/or raw chemicals; *b*) relatively less toxic chemicals, such as sulfonylurea herbicides like nicosulfuron (CASRN 111991-09-4), sulfosulfuron (CASRN 141776-32-1), prosulfuron (CASRN 94125-34-5), and chlorsulfuron (CASRN 64902-72-3), whose reference doses were ranked among the top 15 of the 95 chemicals; and *c*) chemicals subject to relatively rapid transformation in the environment, such as phthalates.

Application 3: Providing Mechanistic Insights into the PROTEX Continuum

Figure 6 shows a breakdown of the estimated emissions by life cycle stages (Figure 6a), receiving compartments (also referred to as the mode of entry; Figure 6b), and the dominant exposure pathways of the general U.S. population to the 95 chemicals (Figure 6c). In general, the relative importance of emission sources, receiving compartments, and exposure pathways were highly dependent on chemical properties and use patterns.

PROTEX-HT predicted that biocides were released mainly during use in agricultural and hygiene settings (yellow bars in Figure 6a). Because the use of biocides in hygiene settings resulted in their occurrence in wastewater, wastewater treatment contributed up to another 40% of the total emission of biocides (navy bars). This was also the case for chemicals in personal care products. Intermediate and/or raw chemicals were predicted to be emitted almost solely during the treatment of industrial waste (navy bars) because they were eventually converted into other components and do not appear in final products. Plasticizers, notably the highly volatile ones, were predicted to be mostly emitted from consumer goods in service (i.e., in-use stock; light blue bars). However, emissions from industrial processes may outweigh those from the in-use phase for less volatile phthalates, for example, di-*n*-octyl phthalate (CASRN 117-84-0) and di(2-ethylhexyl) phthalate (CASRN 117-81-7). Emissions from the in-use stock (light blue bars) were also the most predominant source of construction material additives. Finally, solvents were predicted to be released mainly from industrial activities (pink bars).

The mode of entry of biocides was predicted to be diverse (Figure 6b): More hydrophilic biocides mostly entered surface water (pink bars), whereas more hydrophobic ones mainly ended up in soil (yellow bars). Emissions to surface water (pink bars) arose mostly from wastewater treatment plant effluent. This trend is clearly illustrated by comparing the modes of entry of acephate (CASRN 30560-19-1) and cypermethrin (CASRN 52315-07-8), which have the smallest and highest K_{OW} among all biocides investigated here (Figure S2). Wastewater treatment plants were also predicted to be responsible for the occurrence of intermediate and/or raw chemicals in surface water (Figure 6b). Plasticizers were mostly released into air: The fraction used in consumer goods resulted in their release into indoor air (light blue bars), whereas the fraction released from industrial processes contributed to their occurrence in outdoor air (red bars). Construction material additives also mainly entered outdoor air (red bars). Most chemicals in personal care products were predicted to be applied directly to human skin (navy bars).

Dietary ingestion was predicted to be the predominant pathway for human exposure to intermediate and/or raw chemicals and most biocides (Figure 6c). Dietary ingestion and inhalation were important for solvents. Dermal absorption was predicted to be important for chemicals used in household settings, such as personal care products. Interestingly, the dominant exposure pathway varied among plasticizers: Less volatile phthalates (e.g., di-*n*-octyl phthalate, CASRN 117-84-0; di(2-ethylhexyl) phthalate, CASRN 117-81-7) tended to be ingested via diet and dust, whereas more volatile and water-soluble phthalates (e.g., dimethyl phthalate, CASRN 131-11-3; diethyl phthalate, CASRN 84-66-2) tended to be dermally absorbed and inhaled.

Discussion

Merits of PROTEX-HT

PROTEX-HT succeeded in predicting emissions, fate, exposure, and risk for the 95 case study chemicals. PROTEX-HT reproduced the rates of human exposure to the 95 chemicals inferred from biomonitoring data in the NHANES (Figure 2), as well as contamination in various environmental media and ecological receptors (Table S1). PROTEX-HT predicted the dominant exposure pathways for these chemicals in good agreement with observations reported in the literature. For instance, PROTEX-HT predicted that dietary ingestion contributed to 99% of the overall human exposure to the pesticide chlorpyrifos (CASRN 2921-88-2), higher than the contributions by inhalation (<1%) and dermal absorption (<1%) (Figure 6c). This agreed with the finding that "the major route of chlorpyrifos intake was food ingestion" based on human biomonitoring data (Buck et al. 2001). PROTEX-HT also captured the subtle difference in the dominant exposure pathways between different plasticizers (Figure 6c), which was consistent with biomonitoring evidence that ingestion of food was dominant for di(2-ethylhexyl) phthalate, inhalation was important for dimethyl phthalate, and diethyl phthalate was mostly dermally absorbed (Wormuth et al. 2006).

PROTEX-HT has three major advantages in applications in HT screening-level evaluations of chemicals. First, the modular and integrative nature lends PROTEX-HT great versatility and flexibility in various exposure and risk assessment contexts. PROTEX-HT's components can be run and upgraded either separately or jointly to best fit the purpose of a specific assessment (e.g., indoor vs. outdoor, ecological receptors vs. humans). Application 1 shows that PROTEX-HT's modular nature provides opportunities for the incremental and transparent evaluation of the predictions throughout the production to dose continuum. PROTEX-HT consolidates key components identified in the Aggregate Exposure Pathway framework (Teeguarden et al. 2016; Thayer et al. 2012) and embodies the systems-based approach advocated by the U.S. National Academy of Sciences (NRC 2012). In addition, because PROTEX-HT includes both humans and a broad range of ecological receptors (for details see Text S1, "Description of components in PROTEX-HT" in the
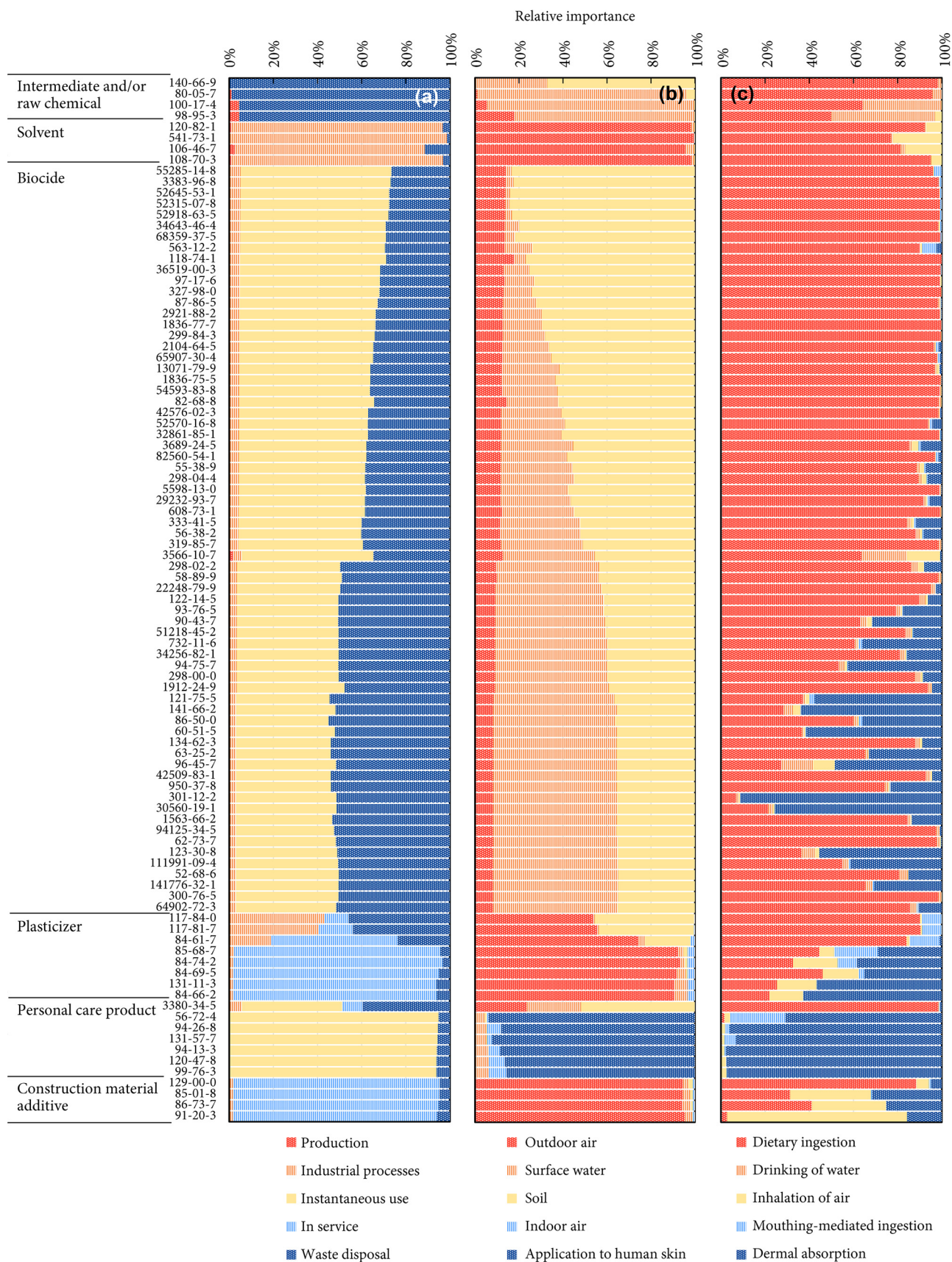
Figure 6. Relative importance of emissions from (a) different life cycle stages to (b) different receiving compartments, and (c) exposure through different routes for 95 synthetic organic chemicals (identified by their Chemical Abstracts Service Registry Numbers) in the United States, categorized by functional category and ranked in the same way as in Figure 5. For the numerical values of the PROTEX-HT predictions, see Excel Table S5. Note: HT, high throughput; PROTEX, PROduction-To-EXposure.

Supplemental Material), it unites screening-level human and ecological assessment end points through a one-health approach. In addition, although we appraised health risks using the external exposure rate in combination with the reference dose that describes *in vivo* systemic toxicity, PROTEX-HT's modular structure enables pairing the whole-body concentrations with *in vitro* bioactivity thresholds for assessments (Turley et al. 2019). Current work indicates that health risk profiles obtained for the same list of chemicals with the two approaches may diverge substantially (Li et al. 2020). PROTEX-HT provides a good opportunity to systematically examine and understand the consistency between these two approaches.

Second, PROTEX-HT's mechanistic basis supports chemical- and population-specific assessments by considering different information on physicochemical properties, use patterns, exposure routes, and toxicity. For instance, Application 3 reveals the dominant emission sources, receiving compartments, and exposure pathways, providing options of different hierarchies for preventing human exposure to harmful chemicals. Eliminating or reducing emissions from the predominant source(s) (Figure 6a) is most proactive at reducing risks; notably, it is inexpensive and simple to implement if a chemical is still at the design or premarket stage. If controlling the predominant source(s) is less practical, then engineering controls could conceivably be implemented to isolate the sources from the major receiving environmental media (Figure 6b; e.g., to ensure chemical use only in closed systems), or to redirect emissions to waste treatment facilities. Finally, if emissions are inevitable, efforts can be made to minimize human exposure from the predominant exposure route(s), such as by removing chemicals from drinking water (Figure 6c). In addition, process-driven models are more informative, transparent, and amenable to testing compared with empirical or statistical approaches usually perceived as a black box. Although this illustrative case focused on the exposure of the general U.S. population, PROTEX-HT's mechanistic nature makes it easy to be reparameterized and tailored to other specific environments and subpopulations that may be more susceptible to certain exposures (e.g., children and workers in developing countries).

Furthermore, PROTEX-HT is input parsimonious. With a reasonable approximation of the level of complexity of the modeled system and the partnership with QSXR techniques, PROTEX-HT can be applied with a minimal number of input parameters: molecular structure and chemical tonnage. This clearly facilitates HT applications for large numbers of chemicals with very limited or no data and saves time and challenges for obtaining input parameters required by the models. It is also possible to replace QSXR predictions with empirical values or professional judgment, if desirable. The application of PROTEX-HT can provide strategic guidance for systematically developing experimental and/or (bio)monitoring programs given that PROTEX-HT prioritizes chemicals of high concern that can be targeted in these programs.

Applicability Domain of PROTEX-HT

PROTEX-HT is designed for synthetic organic chemicals; the use of PROTEX-HT for inorganic substances, metals, and unintentionally generated by-products or transformation products, and so on is not recommended. PROTEX-HT assumes that the technosphere and environment reach and maintain a steady-state, which requires a chemical's time-to-steady-state (i.e., the time at which the total amount of a chemical is close to, e.g., 99% of, the steady-state amount) not to exceed the history of chemical production or the lifetimes of organisms and humans (Li 2020b). Chemicals with a rapid temporal change in tonnage should not be simulated on PROTEX-HT without considering the implications of the steady-state calculations. Further, PROTEX-HT assumes chemicals to be well mixed within homogeneous environmental compartments in a region, making it less suitable for chemicals with extremely rapid loss from a compartment [e.g., chemicals with mass loss more than 25% of the mass in a compartment (see Warren et al. 2009)] and for chemicals that are only released from a single point source. Chemicals outside the above applicability domain will be flagged to allow users to determine and evaluate the reliability of PROTEX-HT predictions. PROTEX-HT's applicability domain is also limited by the combination of applicability domains of the QSXRs implemented in PROTEX-HT. However, PROTEX-HT's modular structure allows users to integrate other QSXRs or empirical data if a chemical falls outside the applicability domains of all QSXRs used here.

Reflection on the Use of Chemical Tonnage in Exposure and Risk Assessments

In Application 1, the comparison between chemical rankings based on exposure and risk indicators (Figure 4; Table S2) reveals that high chemical tonnage does not necessarily result in high exposure rates (Spearman's $\rho = 0.63$), whole-body concentrations (Spearman's $\rho = 0.52$), or health risks (Spearman's $\rho = 0.48$). This finding implies that high production volume (HPV) chemicals are not always of high concern. For example, 4-(1,1,3,3-tetramethylbutyl) phenol (CASRN 140-66-9) is an HPV chemical (ranked 6 in chemical tonnage; Figure 4d) that posed minimal risks to the regional population (ranked 64 in RAF; Figure 4c). This is because only a minuscule fraction of 4-(1,1,3,3-tetramethylbutyl) phenol was released into the environment (Figure 5a). Therefore, it may not be appropriate to use chemical tonnage as a surrogate to prioritize or regulate chemicals. Our finding echoes earlier findings that "chemical tonnages reported to be manufactured or imported played a limited role when determining the potential for overall general human population exposure" (Bonnell et al. 2018).

Nevertheless, it is not feasible to abandon the use of chemical tonnage in regulatory practices given that chemical tonnage is readily measurable, actionable, and verifiable and becomes a part of information requirements in chemical laws and regulations worldwide. One-size-fits-all thresholds, such as the HPV criteria implemented in many countries, may lead to inappropriate over- or underprotection of human and ecological health. An advisable alternative is to set chemical- and population-specific caps for tonnage allowed for production or use in a jurisdiction. Application 2 exemplifies the consideration of variabilities in chemical use, toxicity, and properties to derive chemical- and population-specific maximum allowable tonnage. Furthermore, if the maximum allowable tonnage of each region is summed, one may estimate the total cap of chemical tonnage in the globe, which may be viewed as a planetary boundary for a given substance (Persson et al. 2013).

Compared with chemical tonnage, reliable estimates of emission rates play a more important role in determining exposure and risks. Because emission information is often missing in regulatory practices, chemical tonnages are often used instead as a surrogate to parameterize exposure and risk models. However, as Figure 5A shows, the emission-to-tonnage ratio was always smaller than 1, and the extent to which it deviated from 1 depended closely on chemical properties and use patterns. Using chemical tonnages in models, therefore, overestimated exposure and risks. In this sense, the emission-to-tonnage ratios estimated by the CiP-CAFE module can help parameterize other exposure- and risk-based models.

Future Directions

At present, the quality and availability of chemical tonnage data limit PROTEX-HT's performance. For instance, we used chemical production volumes as a surrogate for tonnages, without considering the potential deviation from actual use amounts owing to import and export, which resulted in a substantial overestimation in human exposure to certain chemicals, such as 4-aminophenol (Figure 2). The 2016 CDR (U.S. EPA 2020) indicates that at least one U.S. manufacturer exported 283 metric tons of 4-aminophenol, accounting for over 50% of the national annual production volume (226–454 metric tons/y; Excel Table S1). However, because most import and export records are labeled as confidential business information, it is challenging to quantify the extent of deviation of the actual use amounts from chemical tonnage. Therefore, more reliable and accurate chemical tonnage data are needed for improving chemical assessments.

PROTEX-HT would benefit from more reliable, accurate parameterization brought about by further advancement of QSXR techniques and databases, as well as other data-mining tools. For instance, the QSUR used here misclassified the functions of certain chemicals and led to biased predictions (Figure 3). Such a bias can be remarkable if the two functions are completely distinct, for example, in the case that a biocide (with negligible consumer contact) is misidentified as a personal care product ingredient (with substantial consumer contact), as demonstrated by examples of nicosulfuron (CASRN 111991-09-4) and sulfosulfuron (CASRN 141776-32-1) in the "Results" section "Evaluation of model performance." However, the bias can also be minor if the two functions share similar use patterns and modes of entry, such as in the case that a construction material additive is misidentified as a colorant [the example of pyrene (CASRN 129-00-0) in the section "Evaluation of model performance"], or a plasticizer is misidentified as a flame retardant. Before such refined QSURs become available, it is recommended to check with empirical records in databases, for example, the Functional Use Database (Isaacs et al. 2016), to confirm the QSUR predictions before use in PROTEX-HT modeling. In addition, the current PROTEX-HT applied the same category-based product information to all chemicals sharing the same functional use category, without considering intra-category variability. Such a limitation can be improved by gathering and curating more detailed market data of products and chemicals.

Future efforts are also encouraged to expand the applicability domain of PROTEX-HT. First, although PROTEX-HT's steady-state predictions represent conservative maximal values and therefore are acceptable (and even recommended) for screening-level assessments (De Bruijn et al. 2002), they may substantially overestimate risks for chemicals persisting in the technosphere and environment, such as persistent organic chemicals used in long-life building materials, because these chemicals may take an unrealistically long time to approach steady-state (Li 2020b). Such situations warrant higher-tiered and time-variant dynamic models, such as PROTEX (Li et al. 2018a, 2018b), as necessary. Second, whereas RAIDAR supports simulations for both neutral and ionogenic organic chemicals (IOCs), RAIDAR-ICE and CiP-CAFE simulate processes relating to the neutral form of IOCs only. Many types of IOCs, such as quaternary ammonium compounds (Li et al. 2020), are also underrepresented in the training sets of the QSXRs used in this work, although PROTEX-HT performed reasonably well for the nine IOCs among the chemicals evaluated herein. PROTEX-HT's modular structure provides great flexibility to incorporate QSXRs designed specifically for IOCs if they become available in the future. Last, PROTEX-HT could be expanded to include more life cycle stages, more environmental compartments, and more exposure routes to increase its versatility. For instance, CiP-CAFE currently does not consider the trace residues of unreacted industrial or raw chemicals in consumer products, such as bisphenol-A remaining in the manufactured polycarbonate plastic products, which led to a substantial underestimation of indoor bisphenol-A contamination in this case study (Table S1). Meanwhile, RAIDAR-ICE does not provide the capability to characterize human exposure to food additives, food preservatives (e.g., parabens; Table S1), or chemicals leached from food packaging materials. The inclusion of new components will complement the current PROTEX-HT modeling capability and support more complete, tailored assessments for chemicals of unique concern.

PROTEX-HT is implemented in a user-friendly online platform, the Exposure And Safety Estimation (EAS-E) Suite (https://www.eas-e-suite.com), to facilitate its application and evaluation by interested parties. The EAS-E Suite platform provides default values for chemical information required to parameterize PROTEX-HT using the existing QSXRs, and databases integrated in the system allow users the option of replacing these parameters with preferred values.

References

Arnot JA, Gouin T, Mackay D. 2005. *Practical Methods for Estimating Environmental Biodegradation Rates. Report for Environment Canada.* CEMN Report No. 2005003. Peterborough, Ontario, Canada: Canadian Environmental Modelling Network, Trent University.

Arnot JA, Brown TN, Wania F. 2014. Estimating screening-level organic chemical half-lives in humans. Environ Sci Technol 48(1):723–730, PMID: 24298879, https://doi.org/10.1021/es4029414.

Arnot JA, Brown TN, Wania F, Breivik K, McLachlan MS. 2012. Prioritizing chemicals and data requirements for screening-level exposure and risk assessment. Environ Health Perspect 120(11):1565–1570, PMID: 23008278, https://doi.org/10.1289/ehp.1205355.

Arnot JA, Mackay D. 2008. Policies for chemical hazard and risk priority setting: can persistence, bioaccumulation, toxicity, and quantity information be combined? Environ Sci Technol 42(13):4648–4654, PMID: 18677986, https://doi.org/10.1021/es800106g.

Bonnell MA, Zidek A, Griffiths A, Gutzman D. 2018. Fate and exposure modeling in regulatory chemical evaluation: new directions from retrospection. Environ Sci Process Impacts 20(1):20–31, PMID: 29271440, https://doi.org/10.1039/c7em00510e.

Breivik K, Arnot JA, Brown TN, McLachlan MS, Wania F. 2012. Screening organic chemicals in commerce for emissions in the context of environmental and human exposure. J Environ Monit 14(8):2028–2037, PMID: 22785348, https://doi.org/10.1039/c2em30259d.

Brown TN, Armitage JM, Arnot JA. 2019. Application of an iterative fragment selection (IFS) method to estimate entropies of fusion and melting points of organic chemicals. Mol Inform 38(8–9):1800160, PMID: 30816634, https://doi.org/10.1002/minf.201800160.

Brown TN, Arnot JA, Wania F. 2012. Iterative fragment selection: a group contribution approach to predicting fish biotransformation half-lives. Environ Sci Technol 46(15):8253–8260, PMID: 22779755, https://doi.org/10.1021/es301182a.

Buck RJ, Özkaynak H, Xue J, Zartarian VG, Hammerstrom K. 2001. Modeled estimates of chlorpyrifos exposure and dose for the Minnesota and Arizona NHEXAS populations. J Expo Anal Environ Epidemiol 11(3):253–268, PMID: 11477522, https://doi.org/10.1038/sj.jea.7500164.

De Bruijn J, Hansen BG, Johansson S, Luotamo M, Munn Sharon J, Musset C, et al. 2002. *Technical Guidance Document on Risk Assessment.* JRC Publication No. JRC23785, EUR 20418 EN. Rome, Italy: European Commission Joint Research Centre.

EC (European Commission). 2007. Regulation (EC) No 1907/2006 of the European Parliament and of the Council Of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH),

establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Off J Eur Union L136:3–280.

ECHA (European Chemicals Agency). 2016a. Chapter R.16: Environmental Exposure Assessment (Version 3.0). In: *Guidance on Information Requirements and Chemical Safety Assessment*. Helsinki, Finland: ECHA.

ECHA. 2016b. *New Approach Methodologies in Regulatory Science. Proceedings of a Scientific Workshop. Helsinki, 19–20 April 2016*. ECHA-16-R-21-EN. Helsinki, Finland: ECHA.

Egeghy PP, Vallero DA, Cohen Hubal EA. 2011. Exposure-based prioritization of chemicals for risk assessment. Environ Sci Policy 14(8):950–964, https://doi.org/10.1016/j.envsci.2011.07.010.

European Chemical Industrial Council. 2021. *The European Chemical Industry: A Vital Part of Europe's Future (2021 Facts and Figures)*. Brussels: European Chemical Industrial Council. https://cefic.org/app/uploads/2021/02/FactsFigures2021_Leaflet_V05.pdf [accessed 24 November 2021].

Fenner K, Scheringer M, MacLeod M, Matthies M, McKone T, Stroebe M, et al. 2005. Comparing estimates of persistence and long-range transport potential among multimedia models. Environ Sci Technol 39(7):1932–1942, PMID: 15871221, https://doi.org/10.1021/es048917b.

Gocht T, Berggren E, Ahr HJ, Cotgreave I, Cronin MTD, Daston G, et al. 2015. The SEURAT-1 approach towards animal free human safety assessment. ALTEX 32(1):9–24, PMID: 25372315, https://doi.org/10.14573/altex.1408041.

Government of Canada. 1999. Canadian Environmental Protection Act, 1999. Canada Gazette Part III. 4 November 1999.

Gramatica P, Cassani S, Chirico N. 2014. QSARINS-chem: Insubria datasets and new QSAR/QSPR models for environmental pollutants in QSARINS. J Comput Chem 35(13):1036–1044, PMID: 24599647, https://doi.org/10.1002/jcc.23576.

Health Canada. 2021. *Science Approach Document – Bioactivity Exposure Ratio: Application in Priority Setting and Risk Assessment*. Ottawa, ON, Canada: Health Canada. https://www.canada.ca/content/dam/eccc/documents/pdf/pded/bioactivity-exposure-ratio/Science-approach-document-bioactivity-exposure-ratio.pdf [accessed 24 November 2021].

Isaacs KK, Goldsmith MR, Egeghy P, Phillips K, Brooks R, Hong T, et al. 2016. Characterization and prediction of chemical functions and weight fractions in consumer products. Toxicol Rep 3:723–732, PMID: 28959598, https://doi.org/10.1016/j.toxrep.2016.08.011.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, et al. 2010. *In vitro* screening of environmental chemicals for targeted testing prioritization: the ToxCast project. Environ Health Perspect 118(4):485–492, PMID: 20368123, https://doi.org/10.1289/ehp.0901392.

Kavlock RJ, Bahadori T, Barton-Maclaren TS, Gwinn MR, Rasenberg M, Thomas RS. 2018. Accelerating the pace of chemical risk assessment. Chem Res Toxicol 31(5):287–290, PMID: 29600706, https://doi.org/10.1021/acs.chemrestox.7b00339.

Li D, Sangion A, Li L. 2020. Evaluating consumer exposure to disinfecting chemicals against coronavirus disease 2019 (COVID-19) and associated health risks. Environ Int 145:106108, PMID: 32927283, https://doi.org/10.1016/j.envint.2020.106108.

Li L, Arnot JA, Wania F. 2018a. Revisiting the contributions of far- and near-field routes to aggregate human exposure to polychlorinated biphenyls (PCBs). Environ Sci Technol 52(12):6974–6984, PMID: 29771504, https://doi.org/10.1021/acs.est.8b00151.

Li L, Arnot JA, Wania F. 2018b. Towards a systematic understanding of the dynamic fate of polychlorinated biphenyls in indoor, urban and rural environments. Environ Int 117:57–68, PMID: 29727753, https://doi.org/10.1016/j.envint.2018.04.038.

Li L, Westgate JN, Hughes L, Zhang X, Givehchi B, Toose L, et al. 2018c. A model for risk-based screening and prioritization of human exposure to chemicals from near-field sources. Environ Sci Technol 52(24):14235–14244, PMID: 30407800, https://doi.org/10.1021/acs.est.8b04059.

Li L, Arnot JA, Wania F. 2019. How are humans exposed to organic chemicals released to indoor air? Environ Sci Technol 53(19):11276–11284, PMID: 31496218, https://doi.org/10.1021/acs.est.9b02036.

Li L, Hughes L, Arnot JA. 2021. Addressing uncertainty in mouthing-mediated ingestion of chemicals on indoor surfaces, objects, and in dust. Environ Int 146:106266, PMID: 33395928, https://doi.org/10.1016/j.envint.2020.106266.

Li L, Wania F. 2016. Tracking chemicals in products around the world: introduction of a dynamic substance flow analysis model and application to PCBs. Environ Int 94:674–686, PMID: 27431909, https://doi.org/10.1016/j.envint.2016.07.005.

Li L, Wania F. 2018. Occurrence of single- and double-peaked emission profiles of synthetic chemicals. Environ Sci Technol 52(8):4684–4693, PMID: 29582660, https://doi.org/10.1021/acs.est.7b06478.

Li L. 2020a. Developing models for tracking the fate of chemicals in products in the total environment. In: *Modeling the Fate of Chemicals in Products*. Singapore: Springer Nature, 27–43.

Li L. 2020b. Introduction: Modeling the fate of chemicals in products in the total environment. In: *Modeling the Fate of Chemicals in Products*. Singapore: Springer Nature, 3–25.

MacLeod M, Fraser AJ, Mackay D. 2002. Evaluating and expressing the propagation of uncertainty in chemical fate and bioaccumulation models. Environ Toxicol Chem 21(4):700–709, PMID: 11951941, https://doi.org/10.1002/etc.5620210403.

Mansouri K, Grulke CM, Judson RS, Williams AJ. 2018. OPERA models for predicting physicochemical properties and environmental fate endpoints. J Cheminform 10(1):10, PMID: 29520515, https://doi.org/10.1186/s13321-018-0263-1.

Murakami S, Oguchi M, Tasaki T, Daigo I, Hashimoto S. 2010. Lifespan of commodities, part I: the creation of a database and its review. J Ind Ecol 14(4):598–612, https://doi.org/10.1111/j.1530-9290.2010.00250.x.

National Academies of Sciences, Engineering, and Medicine. 2017. *Using 21st Century Science to Improve Risk-Related Evaluations*. Washington, DC: National Academies Press.

Nazaroff W, Weschler CJ, Little JC, Cohen Hubal EA. 2012. Intake to production ratio: a measure of exposure intimacy for manufactured chemicals. Environ Health Perspect 120(12):1678–1683, PMID: 23222017, https://doi.org/10.1289/ehp.1204992.

NRC (National Research Council). 2012. *Exposure Science in the 21st Century: A Vision and a Strategy*. Washington, DC: National Academies Press.

OECD (Organisation for Economic Co-operation and Development). 2007. *Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models*. JT03224782. Paris, France: OECD. https://read.oecd.org/10.1787/9789264085442-en?format=pdf [accessed 24 November 2021].

Papa E, Sangion A, Arnot JA, Gramatica P. 2018. Development of human biotransformation QSARs and application for PBT assessment refinement. Food Chem Toxicol 112:535–543, PMID: 28412404, https://doi.org/10.1016/j.fct.2017.04.016.

Papa E, van der Wal L, Arnot JA, Gramatica P. 2014. Metabolic biotransformation half-lives in fish: QSAR modeling and consensus analysis. Sci Total Environ 470–471:1040–1046, PMID: 24239825, https://doi.org/10.1016/j.scitotenv.2013.10.068.

Patlewicz G, Wambaugh JF, Felter SP, Simon TW, Becker RA. 2018. Utilizing threshold of toxicological concern (TTC) with high throughput exposure predictions (HTE) as a risk-based prioritization approach for thousands of chemicals. Comput Toxicol 7:58–67, PMID: 31338483, https://doi.org/10.1016/j.comtox.2018.07.002.

Persson LM, Breitholtz M, Cousins IT, de Wit CA, MacLeod M, McLachlan MS. 2013. Confronting unknown planetary boundary threats from chemical pollution. Environ Sci Technol 47(22):12619–12622, PMID: 23980998, https://doi.org/10.1021/es402501c.

Phillips KA, Wambaugh JF, Grulke CM, Dionisio KL, Isaacs KK. 2017. High-throughput screening of chemicals as functional substitutes using structure-based classification models. Green Chem 19(4):1063–1074, PMID: 30505234, https://doi.org/10.1039/C6GC02744J.

Ring CL, Arnot JA, Bennett DH, Egeghy PP, Fantke P, Huang L, et al. 2019. Consensus modeling of median chemical intake for the U.S. population based on predictions of exposure pathways. Environ Sci Technol 53(2):719–732, PMID: 30516957, https://doi.org/10.1021/acs.est.8b04056.

Rosenbaum RK, Bachmann TM, Gold LS, Huijbregts MA, Jolliet O, Juraske R, et al. 2008. USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. Int J Life Cycle Assess 13(7):532, https://doi.org/10.1007/s11367-008-0038-4.

Rosenbaum RK, Huijbregts MAJ, Henderson AD, Margni M, McKone TE, van de Meent D, et al. 2011. USEtox human exposure and toxicity factors for comparative assessment of toxic emissions in life cycle analysis: sensitivity to key chemical properties. Int J Life Cycle Assess 16(8):710, https://doi.org/10.1007/s11367-011-0316-4.

Sättler D, Schnöder F, Aust N, Ahrens A, Bögi C, Traas T, et al. 2012. Specific environmental release categories—a tool for improving chemical safety assessment in the EC—report of a multi-stakeholder workshop. Integr Environ Assess Manag 8(4):580–585, PMID: 22447453, https://doi.org/10.1002/ieam.1304.

Shin HM, Ernstoff A, Arnot JA, Wetmore BA, Csiszar SA, Fantke P, et al. 2015. Risk-based high-throughput chemical screening and prioritization using exposure models and in vitro bioactivity assays. Environ Sci Technol 49(11):6760–6771, PMID: 25932772, https://doi.org/10.1021/acs.est.5b00498.

Tao M, Li D, Song R, Suh S, Keller AA. 2018. OrganoRelease—a framework for modeling the release of organic chemicals from the use and post-use of consumer products. Environ Pollut 234:751–761, PMID: 29245149, https://doi.org/10.1016/j.envpol.2017.11.058.

Teeguarden JG, Tan YM, Edwards SW, Leonard JA, Anderson KA, Corley RA, et al. 2016. Completing the link between exposure science and toxicology for improved environmental health decision making: the aggregate exposure pathway framework. Environ Sci Technol 50(9):4579–4586, PMID: 26759916, https://doi.org/10.1021/acs.est.5b05311.

Thayer KA, Heindel JJ, Bucher JR, Gallo MA. 2012. Role of environmental chemicals in diabetes and obesity: a National Toxicology Program workshop review.

Environ Health Perspect 120(6):779–789, PMID: 22296744, https://doi.org/10.1289/ehp.1104597.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. 2013. Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect 121(7):756–765, PMID: 23603828, https://doi.org/10.1289/ehp.1205784.

Turley AE, Isaacs KK, Wetmore BA, Karmaus AL, Embry MR, Krishan M. 2019. Incorporating new approach methodologies in toxicity testing and exposure assessment for tiered risk assessment using the RISK21 approach: case studies on food contact chemicals. Food Chem Toxicol 134:110819, PMID: 31545997, https://doi.org/10.1016/j.fct.2019.110819.

U.S. Congress. 2016. Frank R. Lautenberg Chemical Safety for the 21st Century Act. Pub L No. 114–182. Fed Reg 130 Stat. 448. https://www.congress.gov/114/plaws/publ182/PLAW-114publ182.pdf [accessed 24 November 2021].

U.S. EPA (U.S. Environmental Protection Agency). 2012. Estimation Programs Interface (EPI) Suite™ for Microsoft® Windows, v 4.1. Washington, DC: U.S. EPA.

U.S. EPA. 2018. A Working Approach for Identifying Potential Candidate Chemicals for Prioritization. Washington, DC: U.S. EPA. https://www.epa.gov/sites/default/files/2018-09/documents/preprioritization_white_paper_9272018.pdf [accessed 24 November 2021].

U.S. EPA. 2020. Chemical Data Reporting under the Toxic Substances Control Act (2016; Updated May 2020). Washington, DC: U.S. EPA. https://www.epa.gov/chemical-data-reporting/access-cdr-data [accessed 15 November 2020].

Ulrich N, Endo S, Brown TN, Watanabe N, Bronner G, Abraham MH, et al. 2017. UFZ-LSER database v 3.2.1 [Internet]. Leipzig, Germany: Helmholtz Centre for Environmental Research-UFZ. http://www.ufz.de/lserd [accessed 17 August 2020].

Undeman E, McLachlan MS. 2011. Assessing model uncertainty of bioaccumulation models by combining chemical space visualization with a process-based diagnostic approach. Environ Sci Technol 45(19):8429–8436, PMID: 21848272, https://doi.org/10.1021/es2020346.

van de Meent D, de Zwart D, Posthuma L. 2020. Screening-level estimates of environmental release rates, predicted exposures, and toxic pressures of currently used chemicals. Environ Toxicol Chem 39(9):1839–1851, PMID: 32539202, https://doi.org/10.1002/etc.4801.

van de Meent D, Quik J, Traas T. 2014. Identification and preliminary analysis of update needs for EUSES. Contract No. ECHA/2014/253. Bilthoven, Netherlands: National Institute of Public Health and the Environment (RIVM). https://echa.europa.eu/documents/10162/13630/echa_2014_253_euses_report_en.pdf/35a43ff6-4186-4c82-b1fd-8a7742cbfcdf [accessed 24 November 2021].

Vermeire T, Rikken M, Attias L, Boccardi P, Boeije G, Brooke D, et al. 2005. European Union System for the Evaluation of Substances (EUSES): the second version. Chemosphere 59(4):473–485, PMID: 15788170, https://doi.org/10.1016/j.chemosphere.2005.01.062.

Vermeire TG, Jager DT, Bussian B, Devillers J, den Haan K, Hansen B, et al. 1997. European Union System for the Evaluation of Substances (EUSES). Principles and structure. Chemosphere 34(8):1823–1836, PMID: 9114488, https://doi.org/10.1016/s0045-6535(97)00017-9.

Wambaugh JF, Bare JC, Carignan CC, Dionisio KL, Dodson RE, Jolliet O, et al. 2019. New approach methodologies for exposure science. Curr Opin Toxicol 15:76–92, https://doi.org/10.1016/j.cotox.2019.07.001.

Wambaugh JF, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA, et al. 2013. High-throughput models for exposure-based chemical prioritization in the ExpoCast project. Environ Sci Technol 47(15):8479–8488, PMID: 23758710, https://doi.org/10.1021/es400482g.

Wambaugh JF, Wang A, Dionisio KL, Frame A, Egeghy P, Judson R, et al. 2014. High throughput heuristics for prioritizing human exposure to environmental chemicals. Environ Sci Technol 48(21):12760–12767, PMID: 25343693, https://doi.org/10.1021/es503583j.

Wang Z, Walker GW, Muir DCG, Nagatani-Yoshida K. 2020. Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. Environ Sci Technol 54(5):2575–2584, PMID: 31968937, https://doi.org/10.1021/acs.est.9b06379.

Warren CS, Mackay D, Webster E, Arnot JA. 2009. A cautionary note on implications of the well-mixed compartment assumption as applied to mass balance models of chemical fate in flowing systems. Environ Toxicol Chem 28(9):1858–1865, PMID: 19364189, https://doi.org/10.1897/08-569.1.

Weininger D. 1988. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Model 28(1):31–36, https://doi.org/10.1021/ci00057a005.

Wetmore BA, Wambaugh JF, Allen B, Ferguson SS, Sochaski MA, Setzer RW, et al. 2015. Incorporating high-throughput exposure predictions with dosimetry-adjusted in vitro bioactivity to inform chemical toxicity testing. Toxicol Sci 148(1):121–136, PMID: 26251325, https://doi.org/10.1093/toxsci/kfv171.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, et al. 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. Toxicol Sci 125(1):157–174, PMID: 21948869, https://doi.org/10.1093/toxsci/kfr254.

Wignall JA, Muratov E, Sedykh A, Guyton KZ, Tropsha A, Rusyn I, et al. 2018. Conditional toxicity value (CTV) predictor: an in silico approach for generating quantitative risk estimates for chemicals. Environ Health Perspect 126(5):057008, PMID: 29847084, https://doi.org/10.1289/EHP2998.

Wormuth M, Scheringer M, Vollenweider M, Hungerbühler K. 2006. What are the sources of exposure to eight frequently used phthalic acid esters in Europeans? Risk Anal 26(3):803–824, PMID: 16834635, https://doi.org/10.1111/j.1539-6924.2006.00770.x.

Zhang X, Arnot JA, Wania F. 2014. Model for screening-level assessment of near-field human exposure to neutral organic chemicals released indoors. Environ Sci Technol 48(20):12312–12319, PMID: 25264817, https://doi.org/10.1021/es502718k.

Zhang Z, Wang S, Li L. 2021. The role of chemical properties in human exposure to environmental chemicals. Environ Sci Process Impacts, PMID: 34542121, https://doi.org/10.1039/d1em00252j.